# ELT Stack – Proposition 1

Real-World Use Case

Smart City Mobility Insights

Scenario

A mid-sized European city wants to optimize its public transportation system using real-time and historical data. You build an ELT pipeline that ingests data from various sources—bus GPS trackers, bike-sharing APIs, weather feeds, and citizen feedback—and transforms it into actionable insights for city planners.

Goals

* Identify bottlenecks and delays in public transport
* Correlate weather with ridership patterns
* Visualize peak usage zones for bike-sharing
* Monitor citizen satisfaction via sentiment analysis

This use case is rich in data variety, transformation logic, and visualisation potential - perfect for a portfolio.

Synthetic Data Modules Overview

| **Source** | **Sample Fields** |
| --- | --- |
| Bus GPS Trackers | Timestamp, Bus ID, Latitude, Longitude, Speed, Route ID |
| Bike-Sharing APIs | Station ID, Timestamp, Available Bikes, Total Docks, Latitude, Longitude |
| Weather Feeds | Timestamp, Temperature, Precipitation, Wind Speed, Weather Condition |
| Citizen Feedback | Timestamp, Location, Sentiment Score, Comment Text |

| **ELT Stage** | **Tool** |
| --- | --- |
| Generate | Custom Python scripts for synthetic data |
| Extract | Airbyte |
| Load | Airbyte (loads into staging) |
| Staging | DuckDB for lightweight staging |
| Transform | dbt Core (SQL-based transformations, modular and testable) |
| Warehouse | PostgreSQL (or upgrade to Snowflake/BigQuery for scalability) |
| Quality | Great Expectations (data validation and testing) |
| Orchestrate | Prefect (workflow orchestration, retries, scheduling) |
| Visualise | Tableau (dashboards) and Metabase (observability ) |
| Notification | Slack API |
| Monitor | Prometheus + Grafana or Prefect Cloud UI |
| Version Control | Git + GitHub (for dbt models, Prefect flows, documentation) |
| Documentation | dbt Docs + Markdown files + GitHub Pages |

# ELT Stack – Proposition 2

# ELT Stack – Proposition 2